

```
In [1]: import pandas as pd
```

```
In [2]: data=pd.read_csv("/home/placement/Downloads/movies.csv")
```

```
In [3]: data.head(10)
```

```
Out[3]:
```

	srno	movie	year	rating	time
0	1	The Nightmare Before	1993	3.9	4568.0
1	2	The Mummy	1932	3.5	4388.0
2	3	Orphans of the Storm	1921	3.2	9062.0
3	4	The Object of Beauty	1991	2.8	6150.0
4	5	Night Tide	1963	2.8	5126.0
5	6	One Magic Christmas	1985	3.8	5333.0
6	7	Muriel's Wedding	1994	3.5	6323.0
7	8	Mother's Boys	1994	3.4	5733.0
8	9	Nosferatu: Original Version	1929	3.5	5651.0
9	10	Nick of Time	1995	3.4	5333.0

```
In [4]: data.shape
```

```
Out[4]: (49590, 5)
```

In [5]: data.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 49590 entries, 0 to 49589
Data columns (total 5 columns):
#   Column      Non-Null Count  Dtype  
---  -
0   srno        49590 non-null  int64  
1   movie       49590 non-null  object  
2   year        49590 non-null  int64  
3   rating      10814 non-null  float64 
4   time        45836 non-null  float64 
dtypes: float64(2), int64(2), object(1)
memory usage: 1.9+ MB
```

In [6]: data1=data.groupby(['year']).count()

```
In [7]: data1
```

```
Out[7]:
```

	srno	movie	rating	time
year				
1913	3	3	3	3
1914	20	20	5	18
1915	1	1	1	1
1916	1	1	1	1
1918	1	1	1	1
...
2010	5107	5107	1102	4671
2011	5511	5511	1346	4992
2012	4339	4339	1130	3978
2013	981	981	345	901
2014	1	1	1	1

101 rows × 4 columns

```
In [8]: data1.to_csv('movies.csv')
```

In [9]: `data.head()`

Out[9]:

	srno	movie	year	rating	time
0	1	The Nightmare Before	1993	3.9	4568.0
1	2	The Mummy	1932	3.5	4388.0
2	3	Orphans of the Storm	1921	3.2	9062.0
3	4	The Object of Beauty	1991	2.8	6150.0
4	5	Night Tide	1963	2.8	5126.0

In []: